Finally, remembering that $\bar{z} = 0$ and expressing r in terms of product moments about 0 as origin.[1]

$$r_{wz} = \frac{S(wz)/N - \bar{w}\bar{z}}{\sigma_w \sigma_z}$$

and remembering that $\bar{z} = 0$, we may write

$$r_{wz} \frac{S\left\{ w[\Sigma(y) - p\Sigma(x)] \right\}/N}{\sigma_w \sigma_z}.$$

The means of arrays for testing the form of the regression curve of z on w are given by

$$[\Sigma(y) - p\Sigma(x)]/n.$$

J. Arthur Harris

CALCULATION OF THE CORRELATION RATIO

The correlation coefficient r is a good measure of correlation only when the regression is linear. It is necessary, therefore, before placing any reliance upon the computed r to examine the data for linearity of the regression lines. One common test of linearity, Blakeman's, compares the value of $(\eta^2 - r^2)$ with its probable error, η being a correlation ratio. In applying this test it is then necessary to calculate along with r the correlation ratios.

Recently the writer, with a number of assistants, among whom Mr. F. G. Wahlen should be mentioned, had occasion to compute several hundred correlation coefficients, each of which was tested by Blakeman's test. In the progress of the work, in which much use was made of the slide rule, there was developed the tabular form for the calculation of r with the two correlation ratios η_{yx} and η_{xy}, which is the content of this paper. By the addition of two columns to the tabular method of computing r used by Professor H. L. Rietz, the two correlation ratios are easily computed.

This method of calculating the correlation ratios in connection with the calculation of the correlation coefficient is shown in the accompanying numerical illustration. The various columns in connection with the correlation table explain themselves in the headings, with the exception of the columns headed $S_{s'}$ and S_s. The column headed $S_{s'}$ is computed as follows: Let s' be the number of any designated row of the correlation table, say the third row from the top. Then for that row each number is multiplied by the corresponding number in the column headed X. The algebraic sum of these products is $S_{s'}$. For example, for the third row from the top

$$\begin{array}{rl} 9\cdot(-1) = & -9 \\ 26\cdot 0 = & 0 \\ 7\cdot 1 = & 7 \\ 5\cdot 2 = & \underline{10} \\ & 8 = S_{s'} \text{ for } s' = 3. \end{array}$$

The column headed S_s is computed in a similar way.

[1] Harris, "The arithmetic of the product moment method of calculating the coefficient of correlation." *Amer. Nat.*, 44 : 693–699. 1910.

A FORM FOR THE CALCULATION OF THE CORRELATION COEFFICIENT AND THE TWO CORRELATION RATIOS

Monthly pig iron tonnage	Monthly price index 7.50 to 8.00	8.00 to 8.50	8.50 to 9.00	9.00 to 9.50	9.50 to 10.00	f_y	Y	$f_y \cdot Y$	$f_y Y^2$	$S_{s'}$	$S_{s'} \cdot Y$	$S_{s'}^2$	$\frac{S_{s'}^2}{f_y}$
28 to 32				4	4	8	3	24	72	12	36	144	18.00
24 to 28			4	16	1	21	2	42	84	18	36	324	15.43
20 to 24		9	26	7	5	47	1	47	47	8	8	64	1.36
16 to 20	8	21	14	3	1	47	0			−32		1024	21.79
12 to 16	10	8	2	2	1	23	−1	−23	23	−24	24	576	25.04
8 to 12	7	3				10	−2	−20	40	−17	34	289	28.90
Totals						156		70	266		138		110.52

	7.50 to 8.00	8.00 to 8.50	8.50 to 9.00	9.00 to 9.50	9.50 to 10.00	Total
f_x	25	41	46	32	12	156
X	−2	−1	0	1	2	
$f_x \cdot X$	−50	−41		32	24	−35
$f_x \cdot X^2$	100	41		32	48	221
S_s	−24	−5	32	49	18	
$S_s \cdot X$	48	5		49	36	138
S^2_s	576	25	1024	2401	324	
$\frac{S^2_s}{f_x}$	23.04	0.61	22.26	75.03	27.00	147.94

$$\overline{Y} = \frac{70}{156} = .4487$$

$$\sigma_y^2 = \frac{266}{156} - \overline{Y}^2 = 1.5038$$

$$\sigma_y = 1.2263$$

$$\overline{X} = \frac{-35}{156} = -.2244$$

$$\sigma_x^2 = \frac{221}{156} - \overline{X}^2 = 1.3663$$

$$\sigma_x = 1.1689$$

$$\frac{\Sigma X \cdot Y}{N} = \frac{138}{156} = .8846$$

$$r = \frac{1}{\sigma_x \sigma_y}\left\{\frac{\Sigma X \cdot Y}{N} - \overline{X} \cdot \overline{Y}\right\}$$

$$= .687 \pm .029$$

$$\frac{\Sigma \frac{S_{s'}^2}{f_y}}{N} = \frac{110.52}{156} = .7085$$

$$\eta^2_{xy} = \frac{1}{\sigma_x^2}\left\{\frac{\Sigma \frac{S^2_{s'}}{f_y}}{N} - \overline{X}^2\right\}$$

$$\eta_{xy} = .694 \pm .028$$

$$\frac{\Sigma \frac{S^2_s}{f_x}}{N} = \frac{147.94}{156} = .9483$$

$$\eta^2_{yx} = \frac{1}{\sigma_y^2}\left\{\frac{\Sigma \frac{S_s^2}{f_x}}{N} - \overline{Y}^2\right\}$$

$$\eta_{yx} = .705 \pm .027$$

The theory underlying this calculation of the correlation ratios follows from an algebraic transformation of the usual expression for the ratios. The correlation ratio of the y on x, usually represented by η_{yx}, is the ratio of the standard deviation of the means of the arrays of y's (columns), to the standard deviation of the y's. In the notation used in the illustration the mean of each column is given by $\frac{S_s}{f_x}$. The square of the standard deviation of the means of the columns is then

$$\frac{\Sigma f_x\left(\frac{S_s}{f_x} - \overline{Y}\right)^2}{N}$$

where N is the total number of entries in the table. From the above definition of the correlation ratio of y on x we have

$$\eta^2_{yx}=\frac{\Sigma f_x\left(\frac{S_s}{f_x}-\overline{Y}\right)^2}{N\sigma_y^2}.$$

This form can be reduced as follows:

$$\frac{\Sigma f_x\left(\frac{S_s}{f_x}-\overline{Y}\right)^2}{N\sigma_y^2}=\frac{\Sigma\frac{S_s^2}{f_x}}{N\sigma_y^2}-\frac{2\overline{Y}\Sigma S_s}{N\sigma_y^2}+\frac{\overline{Y}^2\Sigma f_x}{N\sigma_y^2}=\frac{\Sigma\frac{S_s^2}{f_x}}{N\sigma_y^2}-\frac{\overline{Y}^2}{\sigma_y^2}$$

remembering that $\frac{\Sigma S_s}{N}=\overline{Y}$ and $\Sigma f_x=N$.

We may then write

$$\eta^2_{yx}=\frac{1}{\sigma_y^2}\left\{\frac{\Sigma\frac{S_s^2}{f_x}}{N}-\overline{Y}^2\right\}$$

and a similar transformation gives

$$\eta^2=_{xy}\frac{1}{\sigma_x^2}\left\{\frac{\Sigma\frac{S_{s'}^2}{f_y}}{N}-\overline{X}^2\right\},$$

both of which are easily obtained from the last columns in the form for computation.

A. R. Crathorne

University of Illinois

NEW SWEDISH PRICE INDEX NUMBERS

The Kommerskollegium of Sweden (Bureau of Commerce) has recently completed an official price investigation undertaken at the close of the war, and the results in the form of a new series of index numbers of Swedish wholesale prices appeared in the official publication of the Kommersiella Meddelanden for May 26, 1922.

The index numbers are built on 160 series of market quotations. Approximately 105 distinct commodities are included, representing raw products and manufactured goods in various stages of elaboration both for producers' and consumers' use. The classification comprises 13 main categories; vegetable food stuffs; animal food stuffs; feed and forage; fertilizers; raw and manufactured products of the iron and metal industries; mortar, brick, cement and glass; lumber; paper and pulp; textile fibres and fabrics; hides, leather, and shoes; rubber; chemical technical products. These main classes are again divided into subordinate commodity groups. Monthly indexes are computed for each series of quotations as well as for each subdivision and general group by taking a